REPORT DOCUMENTATION PAGE

Form Approved
OMB NO. 0704-0188

1. AGENCY USE ONLY (Leave Blank)	2. REPORT DATE 8/13/2002	3. REPORT TYPE AND DATES COVERED April 1, 1998 - March 31, 2001
4. TITLE AND SUBTITLE Final Report: AASERT: Dynamic training of Humans and Tutoring Agents		5. FUNDING NUMBERS N00014-98-1-0435
6. AUTHOR(S) Jordan Pollack		
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Brandeis University 415 South Street Waltham MA 02454		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) Thomas McKenna Office of Naval Research 800 N Quincy St Arlington, VA 22217		10. SPONSORING / MONITORING AGENCY REPORT NUMBER
11. SUPPLEMENTARY NOTES .		
12 a. DISTRIBUTION / AVAILABILITY STATEMENT Unlimited.		12 b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

Our ONR funded research over the past several years has focused on how machine learning systems can continuously improve through the dynamic modification of architecture and the dynamic construction of training environments. Although for different applications we use many different formalisms (neural networks, genetic programs, adaptive dynamical systems), we have focused on a framework for learning in which the environment automatically and incrementally becomes more challenging as the learner progresses. Inspired by "arms race" and sexual selection phenomena in natural evolution, we call this "co-evolutionary learning." It involves a set of learners competing in a game, learning from each other, without a teacher. This AASERT award will take one additional graduate student into an exploration of whether the principles discovered in our machine learning work could be the basis for a new kind of educational technology, where students provide appropriate challenges to each other across the internet, reducing the need for teacher supervision.

14. SUBJECT TERMS Coevolution, Machine Learning, Educational Technology			15. NUMBER OF PAGES
			16. PRICE CODE
17. SECURITY CLASSIFICATION OR REPORT UNCLASSIFIED	18. SECURITY CLASSIFICATION ON THIS PAGE UNCLASSIFIED	19. SECURITY CLASSIFICATION OF ABSTRACT UNCLASSIFIED	20. LIMITATION OF ABSTRACT UL

ASSERT: Dynamic Training of Humans and Tutoring Agents

Final Report

Jordan Pollack
Brandeis University

1. Introduction

It was hoped by many that the advent of computer technology in the classroom would lead to a revolutionary leap in educational outcomes; yet so far, it has not. If children are still arranged with 25 other kids of the same age and economic status, taught by one overwhelmed adult, how will networking the classroom serve to advance education?

Based on our theories of computer learning which derive from many years of ONR sponsored research into the most elementary adaptive systems, such as artificial neural networks and simple reinforcement and evolution systems, and our experience building on-line interactive games like backgammon and TRON, we have begun to develop a new kind of Internet-based educational system. Our software can transport students from the confines of their classrooms, via the generic browsers of their networked workstations, into challenging activity-based communities filled with learners of all ages, genders, and economic status. Learners are placed together because their skill levels and interests are similar, not because they happen to be in the same physical classroom, where competition has been shown to be corrosive (Kohn, 1986).

Our hypothesis is that by tracking user performance, managing the set of available "playmates" for every student, and introducing robotic players at a variety of skill levels, such a community of evolving learners can keep all participants appropriately challenged and motivated to learn.

By perfecting anonymous and indirect interaction we go against the prevailing tide of online communities that allow full communication between participants (e.g. in "chat" rooms and MUD's). There are many reasons for our choice, both scientific and social. Primarily, our learning theory, which is based on a game-theoretic analysis of peer-to-peer competition, **predicts mediocrity when opponents know each other** and can cooperate. Secondarily, we have been experimenting with the provision of **robotic companions** in games, both as the ultimate challenge and to provide easy opponents. Chat would reveal who was a robot!

Unlike an intelligent tutoring system which requires technological support for a detailed model of the ways all students comprehend a subject, our approach only uses technology to support the social construction of the community. Appropriate challenges and opportunities for learning are created by other humans in the learning community. To repeat: this work is not based on strong AI which must understand the student and their knowledge. It is based on weak AI which tracks and matches based on collected data; it is the humans who provide appropriate challenge to each other.

Background

Consider a formal game we call the "teacher's dilemma". One learner, called "teacher" has to provide a sequence of problems or questions to the other learner called "student". The teacher must, through this questioning, learn the extent of the student's knowledge, the boundaries of the student's "zone of proximal development" (Vygotsky 1978), in order to deliver appropriate challenge to the student.

We conceptualize this game as a table. Each time the student gets an answer right, they get a simply payoff, PASS or FAIL. But the "teacher's" payoff is not clear, and is represented by 4 variables. Formalizing the game makes it open to mathematical analysis, and agent simulation.

Student\teacher	Easy	Hard
Right	Pass\Validation	Pass\Joy
Wrong	Fail\Remediation	Fail\Complaint

Figure 1. The "Teachers Dilemma," our formal iterated game, used to explain success and failure of learning, and mediocrity when teacher/student cooperate.

What we discovered is that teachers must be motivated correctly in order to learn about the student; they maximize the utility of seeing their students demonstrate new masteries, or unknown weaknesses, and ignore the pass and fail scores of the student. However, in peer collaboration, and in machine learning through self-play, the "teacher" and "student" can cooperate to share their payoff, even without direct communications or advanced cognition. (Axelrod 1984). A highly motivated teacher, when sharing payoffs with a student, will end up in an equilibrium of "easy question, easy answer". What we have done in our research in machine learning is to identify strategies and environments which ameliorate the collusive practice leading to mediocrity, resulting in longer and more successful applications in our problems of interest. With this grant, we have built learning environments for children on the Internet that turn these machine learning and game-theoretic principles into scientifically sound motivational and scalable educational technology for one-to-one learning.

Evolving game players on the internet

Adjunct to our work in machine learning, we have built successful Internet-based games; and our site has received a surprising number of "hits." First in 1996, to accompany an article in Wired, we put on the Internet a version of our backgammon player (Pollack & Blair, 1998) for humans to play against (visit http://www.demo.cs.brandeis/bkg). This game, though it employs a simple user interface, is still actively played by humans from around the world.

Based on the success of the static backgammon web page, we built a system to allow humans to teach a machine to play a game(Funes, et.al., 1998) We put the simple video game called Tron "light cycles" on the Internet. Humans visit our site and receive in their browser a Java applet that contains a software agent that plays Tron, controlled by a "genetic program" (GP) (Koza, 1992). The agent is one of an evolving population; and the statistics collected on the agents' performance against humans are

used as the fitness function in the co-evolutionary process. Rather than training a single player, we evolved a large community of graded Tron players, each with different tactical characteristics (http://www.demo.cs.brandeis.edu/tron), and our machine learning goal was met.

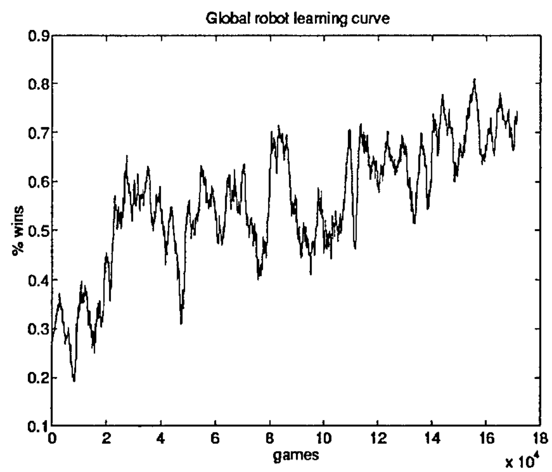


Figure 2. Performance of Tron robots against humanity along the entire experiment (moving average, window size=1000).

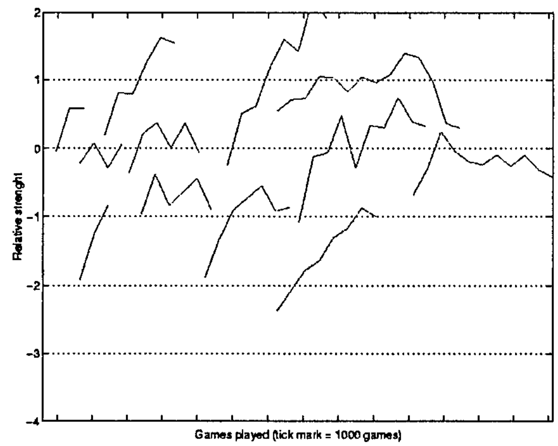


Figure 3. Individual learning: strength curves for the 12 most frequent players (curves start at different x values to avoid overlapping). All users change; nearly all improve in the beginning, but later some of them plateau.

A surprising result of this endeavor was that, despite the clunky interface and antique nature of the Tron game, some humans returned to play thousands of games. Partly, this was due to the novelty of new "life-like" computer players emerging daily. Most video game opponents behave the same way every day, and ours did not. Here is an e-mail excerpt from an avid player who contacted us:

> I can actually see that the best robots now are better than the best robots of yesterday... Yesterday, I could swear that there were sentient beings behind some of the robots... I'm really getting into this game; just played my 1000th game. Now that I know how these robots "think" I can beat a decent number of them without too much mental effort. (Of course, I still have to do a _ton_ of tight maneuvering.)

The Tron experiment has been "on-line" since September 1997 and has had over 5000 visitors, allowing us to observe human behavior and learning. Players new to the system face an initial learning curve. This short-term adaptation is easily measured by averaging the performance of all users over a number of games. Data on each robot — how it does against humans and other robots — is used to create an estimate of each robot's skill, which is constant. We have been able to use these robots as metrics to discover that the humans were actually learning over time.

Related Work

Our work finds itself at the intersection of several areas of human and machine learning, particularly intelligent tutoring systems (ITS) (Clancy, 1986) collaborative environments, and intrinsic motivation. Early seminal work in tutoring systems began with frame-based tutoring systems (Brown & Burton, 1978), moved ahead to support the "learning by doing" view of constructionism (Papert, 1980, Resnick, 1997), and then continued on to focus in areas like constructing rules (ACT) (Anderson, 1982) and modeling student's misconceptions (Van Lehn, 1983, Soloway et al 1981). These ideas were developed into systems (Koedinger and Anderson, 1993, Schank and Cleary, 1995). Student modeling has been aided by statistical techniques and painstaking research (Vanlehn, 1998) as well as artificial intelligence methods like case-based reasoning (Brusilovsky, et al, 1998).

There are many Internet learning community projects: CoVis (Gordon et al, 1996) MariMUSE (Walters and Hughes,1994), KIE (Bell et al., 1995), Belevedere (Suthers, 1997), and MOOSE crossing (Bruckman, 1997) are a few of the more successful examples.

Researchers in human learning and motivation theory have been trying to identify the elements of electronic environments that work to captivate young learners. What makes a child return on his own to a task like a video game, again and again? What is the effect of reward? This has been studied by Lepper and Malone (1987) who separate factors, like fantasy and curiosity, under experimental control. Lepper's own educational software company, (`sparkleinc.com`) highlights factors that have been embedded in their retail edutainment software, such as appropriate challenge, personalization, feedback, recognition, fantasy and choice.

2. Progress

The initial system has been released.

The overall architecture of the system was put in place, and prototype games have been used experimentally with fourth and fifth grade classrooms at a local elementary school. Reactions of students, teachers and parents have been very positive. We were particularly surprised and pleased to note the level of enthusiasm of the students — despite the fact that they are playing a game with rather boring presentation, unadorned with fancy graphics and devoid of any audio. The social aspect of interacting with other humans through the Internet has been enough of a draw to get these students excited, even despite the "drill" like facade of the game.

However, the first small system has several elements of brittleness which must still be addressed. Scientifically, our instruments must capture the right kind of data and be able to control the conditions of user interactions in real-time. The detailed system architecture involves HTML, C, Java and an SQL database. A server mainly acts as a message passer and state maintainer, keeping track of all player clients who are logged into the system and what activity they are engaged in. Small messages sent from the server to the clients will maintain an updated "playground" in each client's browser, illustrating other players coming and going on their own initiative.

So far, we have had 20-30 simultaneous users on a pentium 200. It is clear that as part of continuing work, parts of the software will need to be significantly re-engineered to handle hundreds or thousands of simultaneous games.

The Set of Activities

These server-mediated multi-player environments are **not** the same as board games or video games, but present a whole new experience, where students receive live, slightly lagged, visual and audio feedback of the interaction with peers. Their experience can be dependent on what their playmate does or doesn't do. In deployment, we must pay attention to public aspects of quality, such as the reliability and speed of responses, the aesthetic nature of graphics and sounds, and the interest and frustration level of the users. Initial environments include keyboarding, simple mathematics and geography quiz games in which the students are doing the same task in "timed" trials against each other.

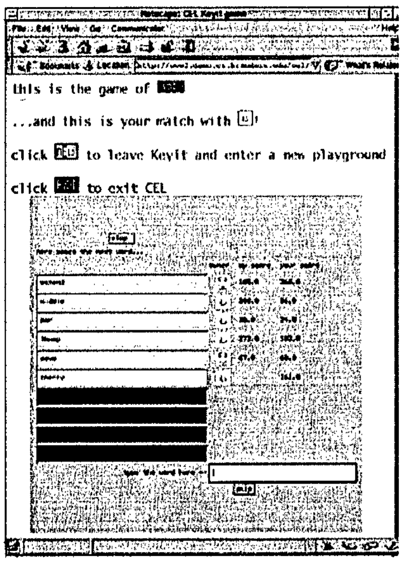
These prototype activities, while seeming like old-fashioned "drill and kill" activities, enabled the initial deployment of server model. We have also introduced games that move beyond the basic skill motif and delve into more complex curricular areas, games in which strategic advancement requires complex integrated cognitive skills rather than primitive skills. For example, we have a two player collaborative "anagrams" game, **MONKEY,** in which players work together to make up words from a set of given letters. We have a "turn-taking" game based on **a realistic engineering model** of LEGO physics in which players work together on collaborative design without communication. It may be one of the first physics-based MUSE's in the world.

In the future, we will be deploying more math/science oriented activities. The **MATHTREE** game uses a grammatical generator of expressions to give each student a target number, and a collection of expressions which they must race to judge as equal to the target. **MathMonkey** gives a target number and a set of elements (numbers, operators). The students must create legal formula equal to the target number.

The ultimate goal of our framework is not only to allow students to **create the challenge** for each other, but moreover to be constrained and motivated to create **appropriate** challenges, just within or just beyond their partner's abilities. This is not transparently easy! In "Wizard of Oz" tests on a two-player spelling bee we are about to release, students were easily able to choose words to stump their opponent. Even if forced to use words they can spell, they can choose "special" rare words they know. In order to constrain the challenge, in the case of a spelling bee, we will give each student a whole sentence, and allow them to choose which word to give their partner. Then we will play each other's sentence for their partner and display text with the chosen word blanked out. Experiments on the reward structure of this game may reveal interesting social psychology issues, similar to the prisoner's dilemma: If I give you a hard word, will you give me a hard word next time? Will different presentations of winning scores affect how students learn and collude?

Results

Our work was originally built as a proof-of-concept to illustrate the idea that we could re-orient the Tron set-up for educational purposes – where human rather than machine learning is the primary goal, and that better pedagogical content than a video game could be delivered. As part of this grant, the system "Community of Evolving Learners" (CEL) was developed into an accessibleand flexible framework for experimenting with and assessing human learning. Forty-five 4th and 5th grade students participated in a pilot study to demonstrate the system in a school setting and to validate CEL's data capture and storage mechanism. The students engaged in two-player keyboarding (typing)



activities. Despite the competitive schema, a post-study survey revealed that all of the students liked the activities.

CEL is a virtual environment implemented on the Internet. Learners engage in interactive educational activities with human partners and/or software agents. The system compiles a database of situated behavior, with which it can track the performance of each participant, building a multi-dimensional model of each learner's changing abilities. The data products collected in CEL support the types of analyses generally performed by researchers in the educational technology field, including: *activity, interaction, improvement, interest* and *coverage.* The graphs below contain data from our pilot study; more data is available in our publications.

Figure 4. KEYIT game screen: Students race each other in typing 10 words, and dynamically see the result of each race.

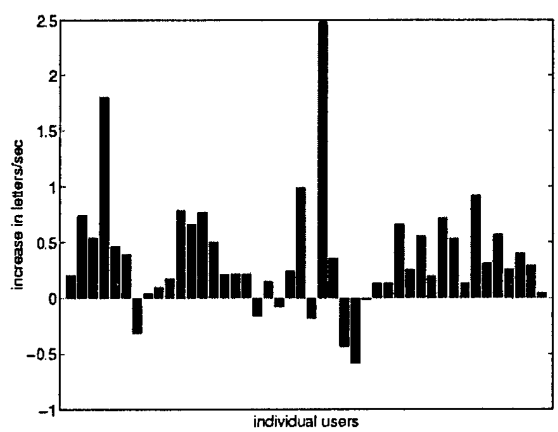


Figure 5. ***Improvement*** **refers to changes in performance. For example, we measured the change in the children's typing speed from the beginning to the end of pilot testing. The figure above illustrates some of the observations. 85% of students improved, showing an increase in typing speed.**

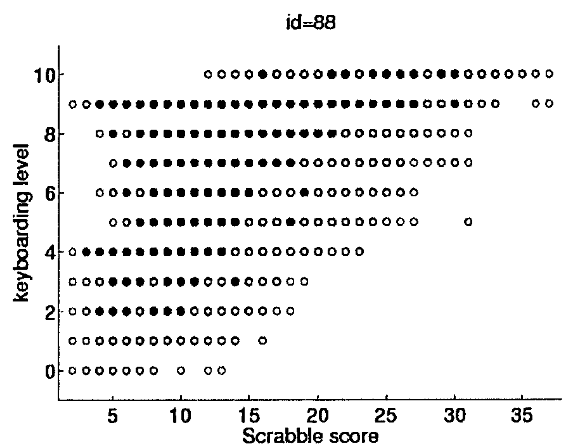


Figure 6. ***Coverage.*** **Our data collection automates feedback from each individual's progress. This graph is from our use of a simple genetic algorithm approach (Holland, 1978 that guided selection of words. We can measure how much of the domain is covered by any individual student (#88) during her participation.**

Publications and Presentations

There was one journal article, several conference proceedings which arose from this ASSERT. Additinally, the graduate student completed her Ph.D., and is now an assistant professor at Columbia University in New York.

- Elizabeth Sklar (2000). **CEL: A Framework for Enabling an Internet Learning Community**. Ph.D. thesis, Department of Computer Science, Brandeis University.

- Sklar, Elizabeth and Pollack, Jordan (2000). A Framework for Enabling an Internet Learning Community. *Educational Technology & Society 3(3)*, Kinshuk, A. Patel and R. Oppermann (eds.).

- Elizabeth Sklar and Jordan B. Pollack (2000). An evolutionary approach to guiding students in an educational game. In Proceedings of the Sixth International Conference on Simulation of Adaptive Behavior (SAB-2000).

- Elizabeth Sklar and Jordan Pollack (2000). Using an evolutionary algorithm to guide problem selection in an online educational game. In Workshop on Evolutionary Computation and Cognitive Science (ECCS 2000). Best student paper award.

- Sklar, E., Blair, A.D. and Pollack, J.B., (2001) Training Intelligent Agents Using Human Data Collected on the Internet, in Agent Engineering, J. Liu, N. Zhong, Y. Y. Tang, and P. Wang (editors), World Scientific Publishing, 2001.

Continued Work

We have successfully sought initial funds from the NSF to continue this line of research. Our goal is to build the infrastructure and prototype games to be able to rapidly test the interaction of variables like competition, collaboration, reward, and anonymity as we assess learning. As part of this AASERT, we were able to develop a prototype, several games, and test it in the classroom. These initial experiments have been limited in our ability to show sustained learning. But the new paradigm leads to an explosion of experimental questions on the basic social psychology of interaction between students and other students as well as robots. In future work we need to establish experimental controls over:

- The availability of specific **activities**;
- The availability of **robot opponents**;
- The availability of playing **modes**, such solitaire, cooperative, or competitive modes, can be varied to each student;
- The **"payoff"** for players who can choose problems for each other can be varied;
- The **level of challenge** offered to each student can be varied to test rates of return and practice;
- The effect of **anonymous** partners and
- Same grade versus multi-grade interactions can be tested.

In addition, with teachers allocating accounts and passwords to their own pupils, they can track progress and engage in controlled studies:

- Entire classes can be enabled to use particular subsets of games in controlled studies prior to standardized testing with equivalent classes who do not have access to the technology.
- Formal computer laboratory use (for entire classes) can be compared to informal individual student use from the classroom and home.

4. Conclusions

A majority of public schools will soon have Internet access in the classroom or computer lab. Many students with these capabilities currently are using them for playing educational games or searching (often pre-screened) sites to supplement traditional research activities. We have been developing machine learners which demonstrate continuous growth through the automatic adaptive nature of environments filled with other learners, and propose to develop and test these environments for humans, initially primary school children. We believe (1) that measuring performance in a game creates a simpler student model than in traditional ITS's, (2) that other humans can create appropriate challenges better than pre-engineered computer programs, and that (3) level of opposition/win-rate is a variable which can be controlled to maximize motivation for students, enticing them to return to our site for more learning. For many content areas, we also can provide a community of graded intelligent software agents to act as assessment tools and to increase human win-rates. Such capabilities could expand students' peer groups beyond their single classrooms, and provide consistent challenges and rewards to all types of learners, independent of factors such as national origin and family income.

Selected Background References

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychology Review, 89.*

Angeline, P. J. and Pollack, J. B. (1994). Competitive environments evolve better solutions for complex tasks. In Forrest, S., editor, *Genetic Algorithms: Proceedings of the Fifth Inter national Conference..*

Axelrod, R. (1984). *The evolution of cooperation. Basic Books, New York..*

Bell, P., Davis, E. A. & Linn, M. C. (1995). The Knowledge Integration Environment: Theory and Design. In J. L. Schnase & E. L. Cunnius (Eds.) *Proceedings of Computer Supported Collaborative Learning (CSCL'95)*, Mahwah, NJ: Lawrence Erlbaum, 14-21. Gordin, D. N., Gomez, L. M., Pea, R. D. & Fishman, B. J. (1996). Using the World Wide Web to Build Learning Communities in K-12.

Brown, J. S. and Burton, R. B. (1978). Diagnostic models for procedural bugs in basic mathematical skills. *Cognitive Science, 2(2).*

Bruckman, A. (1997). *MOOSE Crossing: Construction, Community, and Learning in a networked virtual community for kids. PhD thesis, MIT.*

Brusilovsky, P., Schwartz, E., and Weber, G. (1996). Elm-art: An intelligent tutoring system on the world wide web. In Frasson, C., G., G., and A, L., editors, *Intelligent Tutoring systems, pages 261–269. Springer Verlag.*

Clancey, W. J. (1986). Intelligent tutoring systems: A tutorial survey. Technical Report STAN-CS-87-1174, Stanford University.

Ficici, S. and Pollack, J. (1998). Challenges in coevolutionary learning: Arms-race dynamics, open-endedness, and mediocre stable states. In Adami, Belew, Kitano, and Taylor, editors, *Proceedings of the Sixth International Conference on Artificial Life.*

Funes, P., Sklar, E., Juille, H., and Pollack, J. (1998). Animal-animat coevolution: Using the animal population as fitness function. In Pfeiffer, Blumberg, Wilson, and Meyer, editors, *From Animals to Animats: Fifth International Conference of The Society for Adaptive Behavior. MIT Press.*

Gordin, D. N., Gomez, L. M., Pea, R. D. & Fishman, B. J. (1996). Using the World Wide Web to Build Learning Communities in K-12. *The Journal of Computer-Mediated Communication*, 2 (3)

Koedinger, K. and Anderson, J. (1993). Effective use of intelligent software in high school math classrooms. In *Proceedings of the World Conference on Artificial Intelligence in Education.*

Kohn, A. (1986). *No Contest: The case against competition. Houghton Mifflin.*

Malone, T. W., & Lepper, M. R. (1987). Making learning fun: A taxonomy of intrinsic motivations for learning. In R. E. Snow & M. J. Farr (Eds.). *Aptitude, learning and instruction. Volume 3: Conative and affective process analysis.* Hillsdale, NJ: Lawrence Erlbaum.

Nakabayashi, K., Maruyama, M., Koike, Y., Kato, Y., Touhei, H. and Y. Fukuhara. (1997). Architecture of an Intelligent Tutoring System on the WWW. In *Proc. of the 8th World Conference of the AIED Society (AIED'97).*

Novak, T and Hoffman, D, (1997) New metrics for new media: toward the development of web measurement standards. WWW Journal 2(1), 213-246.

Papert, S. (1980). *Mindstorms: Children, Computers, and Powerful Ideas. BasicBooks.*

Pea, R. (1993). The collaborative visualization project. *Communications of the ACM, 36(5):60–63.*

Pollack, J. and Blair, A. (1998). Coevolution in the successful learning of backgammon strategy. *Machine Learning. to appear.*

Pollack, J., Blair, A., and Land, M. (1996). Coevolution of a backgammon player. In Langton, C., editor, *Proceedings of Artificial Life 5. MIT Press.*

Resnick, M. (1997). *Turtles, Termites, and Traffic Jams : Explorations in Massively Parallel Microworlds. MIT Press, Cambridge.*

Schank, R. and Cleary, C. (1995). *Engines for Education. Lawrence Erlbaum Associates, Hillsdale, NJ.*

Sklar, E., Blair, A., and Pollack, J. (1998). Coevolutionary learning: Machines and humans schooling together. In *Workshop on Current Trends and Applications of Artificial Intelligence in Education: Fourth World Congress on Expert Systems.*

Sklar, E. and Pollack, J. (To appear, 1998). Towards a community of evolving learners. In *Proceedings of International Conference on Learning Sciences.*

Soloway, E. M., Woolf, B., Rubin, E., & Barth, P. (1981). Meno-ii: An intelligent tutoring system for novice programmers. In *Proc 7th Int'l Joint Conf on Artificial Intelligence (IJCAI).*

Suthers, D. and Jones, D. (1997). An Architecture for Intelligent Collaborative Educational Systems. In *Proceedings of the 8th World Conference of the AIED Society (AIED'97).*

VanLehn, K. (1983). Human procedural skill acquisition: Theory, model, and psychological validation. In *Proceedings of the National Conference on AI.*

VanLehn, K. and Martin, J. (1998). Evaluation of an assessment system based on bayesian student modeling. *International Journal of Artificial Intelligence and Education, 8(2)*

Vygotsky, L. S. 1978.*Mind in society: The development of higher psychological processes.* Cambridge, MA: Harvard University Press.

Walters, J. and Hughues, B. (1994). Camp MariMUSE: Linking Elementary and College Students in Virtual Space. *National Educational Computing Conference.*